

2019 Enterprise Connect Verbio



Verbio



AI-POWERED VOICE TECHNOLOGIES

Verbio

In a time when too many businesses are neither delivering memorable customer experiences or achieving operational efficiency, Verbio offers a full range of voice-enabled cognitive solutions supporting intent, context, and preference-driven interactions between humans and machines. Verbio's VoiceAI connections allow enterprises to listen, understand, learn, predict, and engage more naturally with customers and employees alike, putting AI to work to help build context-based, personalized connections. A growing provider with offices in Spain, the USA, Mexico and Brazil, Verbio has processed over three trillion interactions over the past 10 years. Bob Beals, Sales Manager, USA & Canada and Jordi Torres Gomez, General Manager/Managing Partner for Verbio, gave us an in-depth overview.

Why do you consider your mission to use AI to change the way humans and machines interact, to make it less complex, as well as more natural and flexible while allowing for greater personalization?

Verbio has been developing speech technologies and end user solutions for over 20 years. There are two parallel tracks that have always been key: First, the continual optimization of the underlying technology components; the "speech engines" and the servers that host the engines. The improvements in speech engine performance spurred by the emergence of deep neural networks, new linguistic models, and of course the continual increase in CPU clock speeds and chip architecture innovation. From a user interface perspective, we've transitioned from the most constrained grammars (0 - 9, *, #), to saying a predetermined list of specific words, to phrases and short sentences, to now supporting unconstrained conversational speech between a human and machine; all in real-time, without perceptible or unacceptable latency, supporting virtually any language in existence. We can also capture context and intent, all while interacting naturally while riding on a car rental shuttle bus.

Verbio

Second, new devices, new channels, new choices for engagement. Speech user interface design experts have emerged - and become indispensable in developing new applications; Imagining, designing, bringing forth and applying best practices and leveraging known (or anticipated) information about a user - truly approaching an “omnichannel” experience. This is coupled with the ability to deliver a personalized and more natural experience. While we all agree that machines don’t have a brain, the application of AI principals will close the gap and advance what is possible in the realm of human and machine interaction and understanding. Tesler’s Theorem says, “AI is whatever hasn’t been done yet”. At Verbio, we intend to continue to lead in the cognitive space, the human machine interaction space, the creative solutions development space; and continue to do what hasn’t been done before.

Can you explain what you mean by “giving voices to machines?”

“Giving voices to machines” is a reference to Text-to-Speech technology. Historically, when interacting with an IVR, the way to achieve better voice quality was to pay a professional to “pre-record” anything that the machine would ever be anticipated to say. But if a company didn’t have the number, or phrase, or prompt pre-recorded, then that specific function would not work. Every single word had to be pre-recorded either individually or as phrases, sentences, or more. Pre-recorded segments would be concatenated together in real-time to form a complete output element; for example, play a phrase, play a number, play a phrase. The voice output quality was great, but it was a laborious, tedious, constrained, time-consuming, and expensive process. If a business wanted to enhance the solution with a new feature or had many dynamic names or addresses or product names, it would have to pre-record every single element.

Text-to-Speech (TTS) capabilities have been around for many years. The upside of using TTS within a self-service solution was that a business could just feed the IVR a text string, however long, and the TTS engine would generate the audio output. No pre-recording required. The text was fed to the TTS engine, the engine provided speech output, the IVR played the output. This approach presented a lower cost model at scale and faster time to market, but at the expense of quality. Over the years, text-to speech output always had a manufactured sound to it, lacking inflection, intonation, or acceptable adherence to even the simplest linguistic characteristics of the language being used. While functional, it presented a very poor user experience. Verbio now offers TTS technology that can be licensed for use within enterprise solutions, made available as a cloud pay-per-use service for both real-time and offline applications, or used within embedded device applications. The voice quality is nothing like the old “computer generated” TTS, either in demos or actual production consumer-facing solutions. TTS output quality is now approaching actual human voice quality and has reached a point where pre-recording messages is almost over, build and support costs are plummeting, and the user experience has not suffered at all.

To take things a step further, clients are asking us to develop their own, unique Corporate AI Voice as they’ve seen the value in aligning corporate marketing strategies with their service and sales channels. A unique voice which conveys the company’s culture and brand is well worth the investment. Thus, the concept of “giving voices to machines”!

How does your Natural Language Processing capabilities enable businesses to use appropriate grammar, syntax, semantic and statistical models to understand and generate human language?

Verbio

Verbio's Natural Language Processing (NLP) capabilities are foundational to our technology stack. Whether a speech project requires speech to text, text to speech, or voice biometrics capabilities, there are underlying Natural Language Processing services required. Our technology components are used to create solutions that solve business problems and enable more natural human machine interactions.

How does your solution differ from other similar offerings available in the market?

Verbio develops and owns our technology, offering both cloud and on-premise options, enterprise license purchase, license rental, and transactional pricing models. We maintain and invest in an in-house professional services practice including linguists, speech scientists, conversational speech and cognitive user interface design specialists. We make our products and services available at reasonable market price points, and have a reputation for project execution, flexibility, and responsiveness to client requirements – and are known for being easy to work with. We provide technology and solutions that drive business value for enterprises and interaction experiences that make our lives easier.